

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE		5. MONITORING ORGANIZATION REPORT NUMBER(S)	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NPRDC Technical Note 88-12		7a. NAME OF MONITORING ORGANIZATION	
6a. NAME OF PERFORMING ORGANIZATION Navy Personnel Research and Development Center	6b. OFFICE SYMBOL (If applicable) Code 41	7b. ADDRESS (City, State, and ZIP Code)	
6c. ADDRESS (City, State, and ZIP Code) San Diego, CA 92152-6800	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER		
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Technology Headquarters Marine Corps	8b. OFFICE SYMBOL (If applicable) Code 22 Code MA	10. SOURCE OF FUNDING NUMBERS	
8c. ADDRESS (City, State, and ZIP Code) Washington, DC 20350	PROGRAM ELEMENT NO 0602232N 0602131M	PROJECT NO RM3320 CF31P14	TASK NO 04A 02 WORK UNIT ACCESSION NO.
11. TITLE (Include Security Classification) BRAIN ACTIVITY DURING TACTICAL DECISION-MAKING: II. PROBE-EVOKED POTENTIALS AND WORKLOAD (U)			
12. PERSONAL AUTHOR(S) Trejo, L. J., LEWIS, G. W., & Blankenship, M. H.			
13a. TYPE OF REPORT Technical Note	13b. TIME COVERED FROM 85 May to 86 Sep	14. DATE OF REPORT (Year, Month, Day) 1987 December	15. PAGE COUNT 25
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB-GROUP	
05	05	Decision-making, workload, evoked potentials, brain activity, combat systems, performance assessment	
19. ABSTRACT (Continue on reverse if necessary and identify by block number) Evoked potentials (EPs) were recorded from 8 electrodes covering frontal, temporal, parietal, and occipital areas of the brain of 45 Marine Corps subjects during their performance of an air defense radar simulation. Each subject performed the simulation at two levels of workload. EPs were produced by irrelevant, diffuse visual stimuli during performance and during a passive baseline period. Root mean square (rms) amplitude and signal-to-noise ratios were computed for average EPs during each condition. Repeated measures analysis of variance showed that EP signal-to-noise ratios depended on workload, time, and coronal (front-to-back) electrode position, but not on sagittal (left-to-right) electrode position. RMS amplitude and signal-to-noise ratios were as much as 40 percent lower during task performance than during the baseline period. The results show that irrelevant probe EPs provide a measure of brain activity that is sensitive to systematic variations in decision-making workload in a combat system simulation, as predicted by models of mental resource allocation.			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL LEONARD J. TREJO		22b. TELEPHONE (Include Area Code) (619) 553-7981	22c. OFFICE SYMBOL Code 4102

Navy Personnel Research and Development Center

San Diego, CA 92152-6800 TN 88-12 December 1987



LIBRARY
RESEARCH REPORTS DIVISION
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93940

Brain Activity During Tactical Decision-making: II. Probe-evoked Potentials and Workload

Approved for public release; distribution is unlimited.



DEPARTMENT OF THE NAVY
NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER
SAN DIEGO, CALIFORNIA 92152-6800

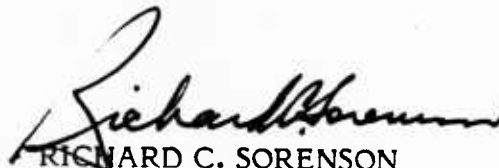
3900
Ser 41/44
13 JAN 1988

From: Commanding Officer, Navy Personnel Research and Development Center

Subj: **BRAIN ACTIVITY DURING TACTICAL DECISION-MAKING: II. PROBE-EVOKED POTENTIALS AND WORKLOAD**

Encl: (1) NPRDC TN 88-12

1. Enclosure (1) is submitted for your information.
2. This technical note is the second in a series of reports examining the feasibility of using neuroelectric signals to predict decision-making of combat system operators under varying workloads. The first report (HFOSL Tech. Note 71-86-6) identified assumptions underlying this approach to the study of decision-making and outlined the experimental hypotheses to be tested. This report provides detailed analyses of the physiological changes in brain activity that occur as workload increases in a combat system simulation.
3. Research described in this report is being performed under the work units Future Technologies--Biopsychometrics (USN Program Element 0602233N, Work Unit RM3320) and Biopsychometric Assessment (USMC Program Element 0602131M, Work Unit CF31P14).
4. Point of contact at the Navy Personnel Research and Development Center is Leonard Trejo, (619) 553-7981 or AUTOVON 553-7981.


RICHARD C. SORENSON
By direction

Distribution:

Deputy Under Secretary of Defense for Research and Engineering (Research and Advanced Technology)
Assistant for Manpower Personnel and Training Research and Development (OP-01B2)
Director, Command Surface Warfare Manpower and Training Requirements (OP-39)
Naval Training Systems Center, Technical Library (5)
Naval Weapons Center, China Lake (Code 3152)
Naval Health Research Center (Code 60) (2)
Naval Research Laboratory (Code 5505) (2)
Naval Aerospace Medical Research Laboratory (Code 031)
Commanding Officer, Naval Health Sciences Education and Training Command, Bethesda, MD
Commanding Officer, Naval Hospital, Portsmouth, VA (Medical Library)
Office of Naval Research (OCNR-114), (OCNR-1142), (OCNR-1142BI), (OCNR-1142CS), (OCNR-1142PS)
Office of Naval Technology (Code 222) (3)

Subj: **BRAIN ACTIVITY DURING TACTICAL DECISION-MAKING: II. PROBE-
EVOKED POTENTIALS AND WORKLOAD**

Headquarters, Marine Corps (Code MA)
Army Research Institute (PERI-SB)
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base, TX
Air Force Medical Research Laboratory (AFMRL/HEG)
Air Force Office of Scientific Research (AFOSR-NL) (2)
Superintendent, Naval Postgraduate School
President, National Defense University (3)
Director of Research, U.S. Naval Academy
Program Manager, Manpower Research and Advisory Service, Smithsonian Institute
Center for Naval Analyses
EEG Systems Laboratory (A. Gevins)
Simon Fraser University (H. Weinberg)
University of California, Los Angeles (J. Beatty)
University of California, San Diego (S. Hillyard, J. Miller) (2)
University of Illinois at Urbana-Champaign (E. Donchin, M. Coles, A. Kramer,
C. Wickens) (4)
University of North Carolina (M. R. Harter)
Washington University (J. Stern)
Defense Technical Information Center (DTIC) (2)

**BRAIN ACTIVITY DURING TACTICAL
DECISION-MAKING: II. PROBE-EVOKED
POTENTIALS AND WORKLOAD**

**Leonard J. Trejo
Gregory W. Lewis
Mark H. Blankenship**

Approved for public release; distribution is unlimited.

**Navy Personnel Research and Development Center
San Diego, California 92152-6800**

SUMMARY

Problem

The demands of many military occupations have the potential for exceeding the capacity of the human to process information, especially during times of great stress, such as those faced by combat system operators. The capacity of the human to perceive, integrate, remember, and use information may be challenged when the individual is flying aircraft, monitoring radar and sonar displays, or operating electronic warfare systems. Exceeding the capacity of the human operator in such situations may impair decision-making and could result in costly tactical errors.

Although much is being done to improve the reliability and effectiveness of combat systems, not enough is being done to improve the system operators. For these reasons, the most unpredictable elements in combat systems are often the human operators themselves. Years of personnel testing using tests such as the Armed Services Vocational Aptitude Battery (ASVAB), have not eliminated this unpredictability. In part, this is because such tests tend to measure what a person *knows* rather than how a person *thinks and reacts*.

This research is driven by the Navy's need for better methods of assessing combat system operators, particularly for predicting the ability of operators to continue to make good decisions under heavy workloads.

Objective

This report, the second in a series of reports concerned with use of neuroelectric signals to predict the decision-making performance of combat system operators, provides detailed analyses of the neuroelectric changes that occur as workload increases in a combat system simulation.

Approach

We presented irrelevant visual stimuli (also called probes) to 45 male U. S. Marine Corps volunteers during a passive baseline period and during their participation in an air defense radar simulation (AIRDEF). Each subject performed the simulation at two levels of workload, which were defined in terms of the number of targets that appeared on the radar display. The probe stimuli were diffuse flashes of light with a duration of 16 milliseconds, presented at irregular intervals.

Under each condition, evoked potentials (EPs), which are physiological measures of brain function resulting from the probe stimuli, were recorded from eight electrodes covering the frontal, temporal, parietal, and occipital areas of the scalp. One electrode was placed above the eye to monitor eye movements. A vertex electrode was the reference for all recordings. Each single EP was first analog-filtered (3 dB bandwidth 0.1-100 Hz), then sampled at 256 Hz, digitized, and stored by a computer. Signal average and signal-to-noise ratio waveforms were computed from six artifact-free EPs for each condition. (Each point in the signal average waveform is the time-indexed average of the six single EPs. Each point in the signal-to-noise ratio waveform is the ratio of the time-indexed average to the time-indexed standard deviation.) Both waveforms were digitally filtered (0.5-25 Hz) and divided into eight adjacent, non-overlapping time windows approximately 50 ms wide that spanned the range between 50 and 450 ms after stimulus onset. The root-mean-square (RMS) value of the waveform was computed in each window for both the signal average (RMS-a) and signal-to-noise ratio (RMS-s) waveforms. These RMS values were used as dependent variables in two repeated measures analyses of variance. Within-subjects factors were *workload*, *coronal* (anterior-posterior) electrode position, *sagittal* (left-right) electrode position, and *time* (window latency).

Results

RMS-s was more sensitive to workload-related effects than RMS-a. RMS-s depended on workload, coronal, and time factors but not on the sagittal factor. Across all sites and time windows, RMS-s was an average of 25 percent lower during AIRDEF performances than during the baseline period. Across time, sagittal electrode positions, and workload levels, RMS-s was greater at the back of the head (parietal and occipital sites) than at the front (temporal and frontal sites). Across the eight time windows, RMS-s peaks occurred between 100-150 ms and between 200-250 ms after stimulus onset.

An interaction between coronal and time factors showed that RMS-s between 100-150 ms was larger at the occipital sites than at all other sites. This large occipital RMS-s peak appeared to be insensitive to workload. However, workload-sensitive peaks were identified in a significant three-way interaction of the workload, coronal, and time factors. Frontal 100-150 ms and 250-350 ms peaks and parietal/occipital 200-250 ms peaks were reduced from 33 to 47 percent during AIRDEF performance as compared with those occurring during the baseline condition. A latency increase in the frontal RMS-s peak near 300 ms was also suggested by this interaction.

Results for RMS-a were similar to those for RMS-s; however, the workload factor was not significant for RMS-a. Other differences between the two measures were consistent with the hypothesis that for small numbers of single EPs, a signal-to-noise ratio measure, such as RMS-s, is influenced less by large random potentials than is a signal average measure, such as RMS-a.

The results are discussed in relation to a *resource allocation model* of cognitive processing that predicts smaller probe EPs under high workload conditions than under baseline or low workload conditions. Implications of the results for the development of EP predictors of decision-making performance by combat system operators are discussed.

CONTENTS

	Page
INTRODUCTION	1
METHODS	3
Subjects	3
Task	3
Instructions to Subjects	4
Probe Stimulus	4
EP Recording	4
Artifact Rejection	5
Selection of EPs for Analysis	6
Signal Processing	6
Analyses	8
RESULTS	9
Synopsis	9
Effects of Workload	11
Differences Related to Electrode Placement	13
Differences related to Post-stimulus Time	14
Interaction of Time and Electrode Placement	14
Interaction of Workload, Time, and Electrode Placement	15
DISCUSSION	18
CONCLUSIONS	20
RECOMMENDATIONS	21
REFERENCES	23

LIST OF TABLES

	Page
1. Summary of Biographical Data	3
2. Time Windows for Analyses of EP Waveforms	9
3. Analyses of Variance of Evoked Potential RMS-s and RMS-a	11
4. Site- and Time-specific Comparisons of Baseline with Active (AIRDEF) EP Measures	17

LIST OF FIGURES

1. Photograph of the air defense simulation displayed on a Masscomp 5500 computer graphics terminal	5
2. AWAVES (average EP) and SNWAVES (signal-to-noise ratio) of one subject shown for eight recording sites across experimental conditions	7
3. AWAVES for 12 subjects based on six single, visual-probe EPs recorded over the left occipital area	8
4. A schematic view of the human head from above	10
5. Mean RMS-s collapsed across eight time windows and plotted as a function of workload	12
6. Mean RMS-s and RMS-a collapsed across eight time windows and plotted as a function of coronal electrode position	13
7. Mean RMS-s and RMS-a collapsed across eight recording sites and plotted as a function of time after stimulus onset	14
8. Mean RMS-s collapsed across homologous recording sites in the same coronal plane and plotted as a function of time after stimulus onset	15
9. Mean RMS-s and RMS-a collapsed across homologous recording sites in the same coronal plane and plotted as a function of time after stimulus onset and separated by workload level	16

INTRODUCTION

The Navy depends heavily on written testing to evaluate personnel. The results of such testing can predict academic performance reasonably well, but are less effective in predicting on-job performance. There is a need for new kinds of testing procedures that will supplement the information derived from written tests and provide an improved understanding and a more complete assessment of the unique capability of each individual.

One of the newest and most promising techniques for personnel assessment is direct measurement of brain activity, done with or without engagement of the subject in a task (e.g., Lewis, 1983). For example, visual evoked potentials produced during passive observation of a flashing light have shown promise for predicting sonar operator (Lewis & Rimland, 1980) and aviator performance (Lewis & Rimland, 1979). Research on brain activity recorded with the subject engaged in a task has been related to performance in tactical jet aircraft (Karis, Coles, & Donchin, 1984), display-monitoring performance (Israel, Wickens, Chesney, & Donchin, 1980), and complex perceptual-motor task performance (Kramer, Wickens, & Donchin, 1983).

The emphasis of the present research is on using brain activity to assess and ultimately to predict decision-making performance of combat system operators. A previous report (Trejo, 1986) discussed the background for this research, including assumptions, hypotheses, and experimental design. An air defense radar simulation (AIRDEF) (Kelly, Greitzer & Hershman, 1981) was adapted as a tool for assessing brain activity and behavior during decision-making under various workloads. Some of this background is summarized and expanded here.

Three theoretical constructs together form a chain that links recordings of brain activity to assessment and prediction of decision-making. The first construct in the chain is a *mapping hypothesis* that relates behavior to brain electrical activity. This hypothesis states that brain activity is the causal agent underlying both behavior and neuroelectric signals (Callaway, 1975; Regan, 1972; Trejo, 1986). Thus, each distinct behavior, including "mental" behavior, corresponds (probabilistically) to a distinct set of neuroelectric signals at the scalp. Regan (1972, p. 31) points out that because neuroelectric signals can relate either to stimulus properties or to properties of sensations produced by stimuli, we cannot tell whether brain activity produces sensations or the sensations produce brain activity. However, a causal relationship between brain activity and behavior may be inferred when specific neuroelectric signals reliably precede certain behavior. For example, recent experiments have shown that different patterns of correlations between signals from various brain regions precede accurate and inaccurate performance in a perceptual-motor task (Gevins, 1986). Whether a causal relationship may be inferred or not, neuroelectric signals may still act as a code or a sign of sensations or other mental events, and, for our purposes, a cause-and-effect relationship need not necessarily be determined. To be useful, the signals only need to indicate the likelihood of a mental event or observable behavior.

The second link is a *resource allocation model* of brain processing (Moray, 1967; Norman & Bobrow, 1975; Taylor, Lindsay, & Forbes, 1967; Wickens, 1984) that states that the total capacity of the brain, which is limited (Broadbent, 1958), can be divided and allocated in varying proportions to separate tasks. To illustrate this, consider a person engaged in a complex task, such as driving a car. Clearly, under ideal driving conditions, the driver would be able to operate the car and admire some of the scenery. However, under more demanding conditions, such as in heavy traffic, the driver is likely to miss many details of the scenery even though they are still present in his field of view. In this scenario the processes associated with driving are critical, whereas those associated with admiring the scenery are non-critical. This distinction arises because safe maneuvering of the car assumes a higher priority than viewing scenery. In terms of the resource allocation model, when the performer of a complex task allocates more brain resources to critical tasks in order to meet increased processing demands (workload), then fewer resources can be allocated to non-critical tasks. Consequently, in order to hold performance constant on critical sub-tasks, performance on non-critical sub-tasks may suffer.

Together, the resource allocation model and the brain-behavior mapping hypothesis predict that changes in the neuroelectric correlates of a non-critical task in a combat system simulation will reflect changes in the allocation of resources to critical tasks. To illustrate this, consider again the driving example. If a component of the scenery in the field of view of the driver were a flashing neon sign, then recordings of brain activity would show visual evoked potentials synchronized with the onset of

the flashes. What the resource allocation model predicts is that more resources will be allocated to driving and fewer resources to attending to the flashing sign under conditions of heavy traffic than under ideal driving conditions. Perception of critical stimuli, such as other vehicles, is facilitated by strategic allocation of brain resources, while perception of irrelevant stimuli, such as the neon sign, is suppressed. The brain-behavior mapping hypothesis translates these differences in allocation of resources into differences in the evoked potentials synchronized with the flashing sign. The nature of these evoked potential differences is not necessarily specified by either the resource allocation model or the brain-behavior mapping hypothesis. Instead, these differences must be inferred from experimental data.

Typically, workload-related changes in both the amplitude and latency of evoked potential components, such as P300, have been demonstrated.¹ In one study with direct relevance to the present research (Wickens, 1984), measures of P300 latency and amplitude were related to resource allocation. When combined with measures from a behavioral test battery, these evoked potential measures accounted for 18 to 41 percent of additional variance in the subsequent performance of a complex combat system simulation. Thus, in our driving example, we may expect to find that the evoked potentials associated with the flashing sign would be larger and faster under ideal driving conditions than in heavy traffic, even when the sign occupies the same field of view. Such evoked potential changes are the basis of the probe-evoked potential technique (Papanicolaou & Johnstone, 1984) and of our use of AIRDEF, which combines a complex radar simulation (critical task) with an irrelevant visual stimulus (non-critical task).

The direction and magnitude of evoked potential changes under varying workloads may not be the same in all individuals. Rather, these changes could be determined by an individual's total processing capacity and by his or her resource allocation strategy. Thus, individual patterns of evoked potential change under simulated workloads may serve as predictors for these individuals of future performance under real workloads.

The third and final link in the chain between decision-making assessment/prediction and neuroelectric signals is what we shall call the *principle of covert behavioral compensation*. What we have in mind here is an explanation for the demonstrated ability of evoked potential measures to account for variance in performance over and above the variance accounted for by behavioral measures. We postulate that when processing capacity is in great demand, subjects may perform covert (unobserved or unmeasured) behavior that counteracts the increased workload and prevents a performance decrement. In a complex task, it is virtually impossible to observe and measure all of the behavior involved in the performance. Some covert behavior is measurable, but is not typically measured. Examples include muscle tension, posture, breathing, sub-vocalization, eye movements, pupil size, heart rate, blood pressure, and so on. Other covert behaviors are not directly measurable. Mental exercises such as counting, rehearsing, and concentrating are strategies that a subject may employ to improve performance but that cannot be observed. It follows from the brain-behavior mapping hypothesis that correlates of covert compensatory behavior will exist in concurrently recorded neuroelectric signals. Thus, the neuroelectric signals may predict or account for variance in performance that is not accounted for by behavioral measures alone.

We now report an analysis of workload-related changes in the neuroelectric correlates of a visual probe stimulus recorded during the performance of AIRDEF by a group of U. S. Marines.

¹Evoked potential components are wavelets or deflections that are identified by specific latency and polarity criteria. For example, P300 refers to a positive wave with a latency of about 300 milliseconds.

METHODS

Subjects

Complete evoked potential (EP) data were recorded from 45 male volunteers from the U. S. Marine Corps. Each subject received a brief description of the research objectives and techniques, filled in a brief biographical questionnaire, and signed a voluntary Privacy Act notice and legal consent form. The questionnaire results, summarized in Table 1 showed that the group consisted of mostly right-handed Caucasians in their early twenties. Furthermore, most subjects reported that they were in good health, had not been taking any medication and had good vision and hearing. Self-assessment of arousal was performed on a four-point scale, with "tired" and "alert" at the extremes. About half the subjects reported being either tired or drowsy. More than half of the subjects smoked cigarettes or chewed tobacco. No attempt to relate biographical data to EP measures will be made in this report.

Task

Our version of the air defense simulation (AIRDEF) adhered to the original specifications of Kelly et al. (1981) with two exceptions. First, in our version, the display was color-coded (blue radar sweep and range indicators, amber targets, weapons, track numbers, and launch indicator). Second, the time required for one complete revolution of the radar sweep was 6.5 seconds. In the Kelly et al. version of AIRDEF, the display was monochromatic and the sweep revolved once every 11.7 seconds. A more detailed description of our version of AIRDEF appears in Trejo (1986).

Each experimental run consisted of three conditions: one passive baseline condition, and two active AIRDEF engagements. Each subject participated in the baseline condition first, in which 10 visual probe stimuli were presented, but no simulation activity occurred. Except for the intermittent appearance of the probe stimuli, the display remained blank. The subject was instructed to observe the display, but no visual fixation point was provided. Next, each subject performed the AIRDEF simulation at *level 1*, in which workload, as measured by target frequency, was 4.5 targets per minute. Finally, each subject performed the simulation at *level 2*, in which workload was 9 targets per minute.

Table 1

Summary of Biographical Data

Age (years)		Height (inches)		Weight (pounds)	
Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
21.8	1.7	69.96	2.48	164.96	19.21
Race		Other 2%	Handedness		
Caucasian 85%	Black 13%		Left 4%	Right 92%	Ambidextrous 4%
Eye Color			Hair Color		
Brown 47%	Blue 31%	Green 22%	Brown 67%	Black 20%	Blond 13%
Arousal			Languages Spoken		
Tired 27%	Drowsy 18%	Awake 35%	Alert 20%	English Only 93%	English and Other 7%
Tobacco Use			Medication in Last 24 Hours		
Smoke or Chew 60%	No Tobacco Use 40%		No 93%	Yes 7%	
Wear Glasses/Contact Lenses			Difficulty Hearing		
No 76%	Yes 24%		No 93%	Yes 7%	

The total engagement time for each level was 4 minutes. During each level 20 probe stimuli were presented. Following each of the two active engagements, the experimenter showed each subject his performance measures and answered any of his questions about the simulation.

Each subject was assigned to one of two groups. Subjects in the *typed* command group (N = 23) launched weapons by typing the track numbers on a keyboard, as in the Kelly et al. version of AIRDEF. Subjects in the *voice* command group (N = 25) launched weapons by telling an experimenter which track numbers to fire on; the experimenter, in turn, typed the numbers on the keyboard. Due to their need to enter track numbers on the keyboard, subjects in the typed command group were expected to look away from the display more often than those in the voice command group.

Instructions to Subjects

Using a labeled reproduction of the AIRDEF display (Figure 1), an experimenter explained the simulation to each subject. The experimenter did not allow the engagement to begin until he felt sure that the subject understood the meaning of the display symbols, the method of firing weapons, and the rewards or penalties associated with the various possible outcomes of his actions (hits, kills, splashes, in-flight launches). The experimenter also explained the strategy for maximizing the performance score. This involved telling each subject how far to allow incoming missiles of various speeds to travel before firing at them in order to obtain a maximum-range kill. Subjects were further instructed not to move excessively or to talk unnecessarily during the task engagements.

Subjects were told that a visual stimulus for producing EPs would appear intermittently during baseline and active engagements. They were also told that the probe had no relevance to the task and that it should not be attended to or counted.

Probe Stimulus

The probe stimulus was a filled neutral gray rectangle that appeared briefly on the same 13-in. color monitor (Ikegami CDA series)² used for the AIRDEF display. (The rectangle completely covered the addressable area of the display.) When viewed from the typical operating distance (45 cm), the stimulus subtended 37° visual angle vertically and 49° horizontally. The stimulus persisted for 16 ms, and had a luminance of about 3 footlamberts. This stimulus was a *background* flash; it did not interfere with the observation of the task on the screen. In all conditions, the stimuli appeared in a pseudorandom time sequence. The average interval between stimuli was 11 seconds with a range of 6 to 16 seconds.

EP Recording

The recording system consisted of a standard array of tin electrodes embedded in a nylon cap (Electro-Cap International, Inc.), a set of independent amplifiers (Grass Model 12A, Neurodata Acquisition System) and a computer (Masscomp 5500) programmed to digitize and record neuroelectric signals. The signals were band-pass filtered (analog, 3 dB corner frequencies at 0.1 & 100 Hz), amplified (20,000 times), sampled at 256 Hz, and digitized with an overall precision of 0.1 μ V.

The basic physiological data were single-epoch visual EPs. For a period of 500 ms before and 500 ms after each probe stimulus (1 second total), the computer stored the potentials recorded from an array of electrodes on the subject's scalp. The array consisted of eight sites: F3 & F4, T3 & T4, P3 & P4, O1 & O2, which covered the left and right frontal, temporal, parietal, and occipital regions of the head (Jasper, 1958). We also recorded from a site above the orbit of the right eye, Fp2 (right frontal pole), to monitor eye blinks and large eye movements. All recordings were bipolar with the reference electrode at Cz, the vertex (intersection of the mid-sagittal and inter-aural lines). Subject ground was at a point on the midline, approximately 3 cm anterior to the frontal sites.

²Identification of the equipment here and elsewhere in the report is for documentation only and does not imply endorsement.

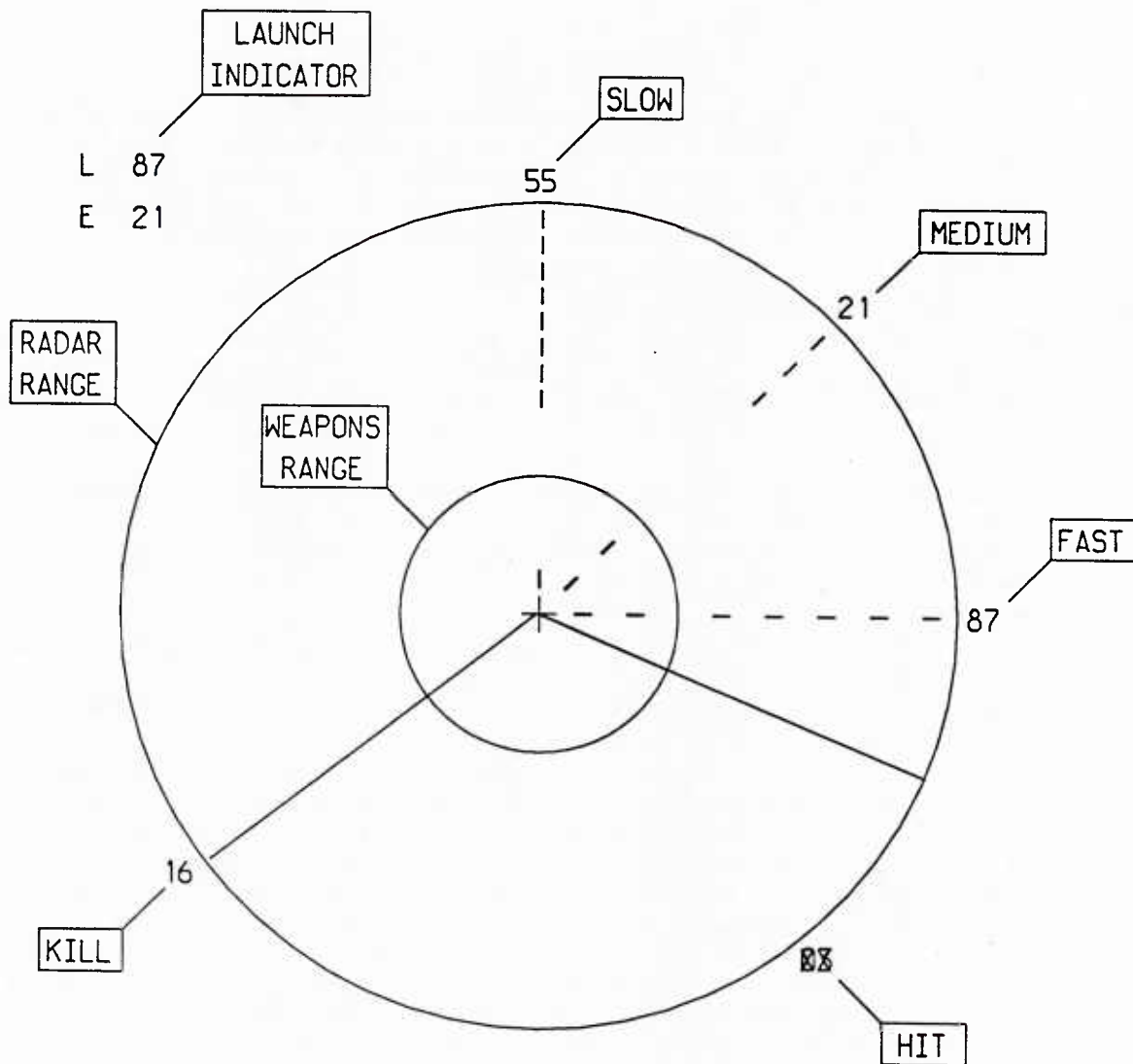


Figure 1. Photograph of the air defense simulation displayed on a Masscomp 5500 computer graphics terminal. The cross in the center shows the location of the subject's ship and the inner circle marks the maximum weapons range. The outer circle marks the maximum detection range for the radar. Numbers along the outer circle identify incoming targets. "Blips" (short line segments that appear silently) represent present and past positions of either incoming targets or outgoing weapons. On the display, these are color-coded with orange and green, respectively. The letters in the upper left corner of the display provide feedback to the operator after each attempted weapon firing. *L* indicates a successful launch attempt on the adjacent track number and *E* indicates an unsuccessful launch attempt, which occurs if an invalid track number is entered or if the requested track already has a weapon in flight. For example "E 21" in the figure means that the weapon fired on track 21 was either too early (splash) or that an inflight launch attempt was made on track 21. For further details, see Kelly et al. (1981) and Trejo (1986).

Artifact Rejection

In order to remove eye blink or eye movement artifacts, we defined a criterion for rejection based on the signals recorded at site Fp2, near the right eye. The criterion was any transient signal with a baseline-to-peak amplitude greater than 50 μ V during the 500 ms post-stimulus period.

In four subjects, the video monitor produced a regular electrical artifact at stimulus onset that resolved to baseline during the first 100 ms of the post-stimulus period. We removed this artifact by computing the average wave shape during the first 100 ms across the eight scalp sites and then subtracting this average wave from the EP at each site.

Selection of EPs for Analysis

Since the number of artifact-free epochs varied between subjects, we developed three criteria for retaining artifact-free epochs for analysis. First, the number had to be equal for all three conditions (baseline, level 1, level 2). Second, the number had to be equal across all subjects. Third, the number had to be one that would retain a majority of our sample of subjects for analysis. We found that 30 subjects, or 65 percent, had at least six artifact-free EPs in all three recording conditions.

Because workload (target frequency) in AIRDEF rises at the beginning of the engagement and falls near the end, EPs recorded near the beginning or the end may not show workload-related effects. To increase the sensitivity of our EP measures to workload, we excluded the earliest and latest EPs from analysis whenever more than six artifact-free EPs were available. Therefore, the six EPs retained were generally from the middle of the AIRDEF engagement where workload was nearly constant.

Signal Processing

From the six selected EPs for each subject and condition, we computed the average EP waveform. For brevity, we will refer to this measure as the *AWAVE*. Each point in a single EP represents the voltage measured between an electrode and the reference at a certain time relative to the onset of the stimulus. Each point in the *AWAVE* is the arithmetic mean of the corresponding time-indexed points in the six single EPs. Because it is a signal average,³ the *AWAVE* emphasizes evoked potentials and de-emphasizes noise. It is reliable for large numbers of single EPs. However, for small numbers of EPs, a single EP containing a large amount of noise can unduly influence the resulting *AWAVE*. Such an *AWAVE* may be indistinguishable from one representing the average of several EPs, each containing less noise and a small, but consistent, evoked potential.

To guard against spurious results that can be caused by noise we developed a signal-to-noise amplitude ratio measure of the average EP, which we shall refer to as the *SNWAVE*. Each time point in the *SNWAVE* is the ratio of the arithmetic mean to the unbiased standard deviation of the six corresponding points in the six single EPs. Like the *AWAVE*, this is a composite measure that emphasizes evoked potentials and de-emphasizes noise. However, unlike the *AWAVE*, the magnitude at any point in the *SNWAVE* is inversely proportional to the variability at corresponding points in its constituent single epochs. Thus, the contaminating effects of large noisy deflections are attenuated. Our definition of the *SNWAVE* is similar to that of another signal-to-noise ratio measure that has been successfully applied to neuroelectric recordings (John et al., 1983). Because both the mean and the standard deviation have the same units (μV), the values of points in the *SNWAVE* are dimensionless.

Both the *AWAVEs* and *SNWAVEs* were de-trended (mean and linear slope removed) and digitally filtered (windowed finite impulse response filter, 127 coefficients, corner frequencies at 0.1 and 25 Hz, Hamming window). Figure 2 shows the *AWAVEs* (2A.) and *SNWAVEs* (2B.) for all three conditions recorded at eight different sites in one subject. The adjacent traces in the 12 sections of each graph are from the left and right recording sites. In both the *AWAVE* and *SNWAVE* records, a large positive deflection occurs between 100 and 200 ms after stimulus onset (time zero on the ordinate) at all sites in the baseline condition and at the parietal and occipital sites in levels 1 and 2. The oscillatory activity at the temporal sites in levels 1 and 2 is often the result of muscle artifact. Note how this oscillatory activity is attenuated in the *SNWAVEs*. Note also that the evoked components at the occipital and parietal sites seen in the *AWAVEs* are preserved or enhanced in the *SNWAVEs*.

Not every subject displayed the same EP morphology. Figure 3 shows that there was a wide range of individual differences in the amplitude and latency (time to peak) of the positive deflection near 100 ms and also among deflections at higher latencies. Because our subjects did not all display

³By "signal averaging" we mean synchronous averaging, which has also been called the method of computed averaged transients. The method assumes that the signal to be measured is weak and is embedded in noise, but can be triggered repetitively. Each measurement of the signal is synchronized with the triggering event. The set of measurements is then averaged synchronously with respect to the triggering event to compute a better estimate of the signal than that provided by any single measurement alone. This process enhances the signal-to-noise ratio of the signal estimate by a factor equal to the number of measurements that are averaged. In our experiments, the triggering event is the visual probe stimulus.

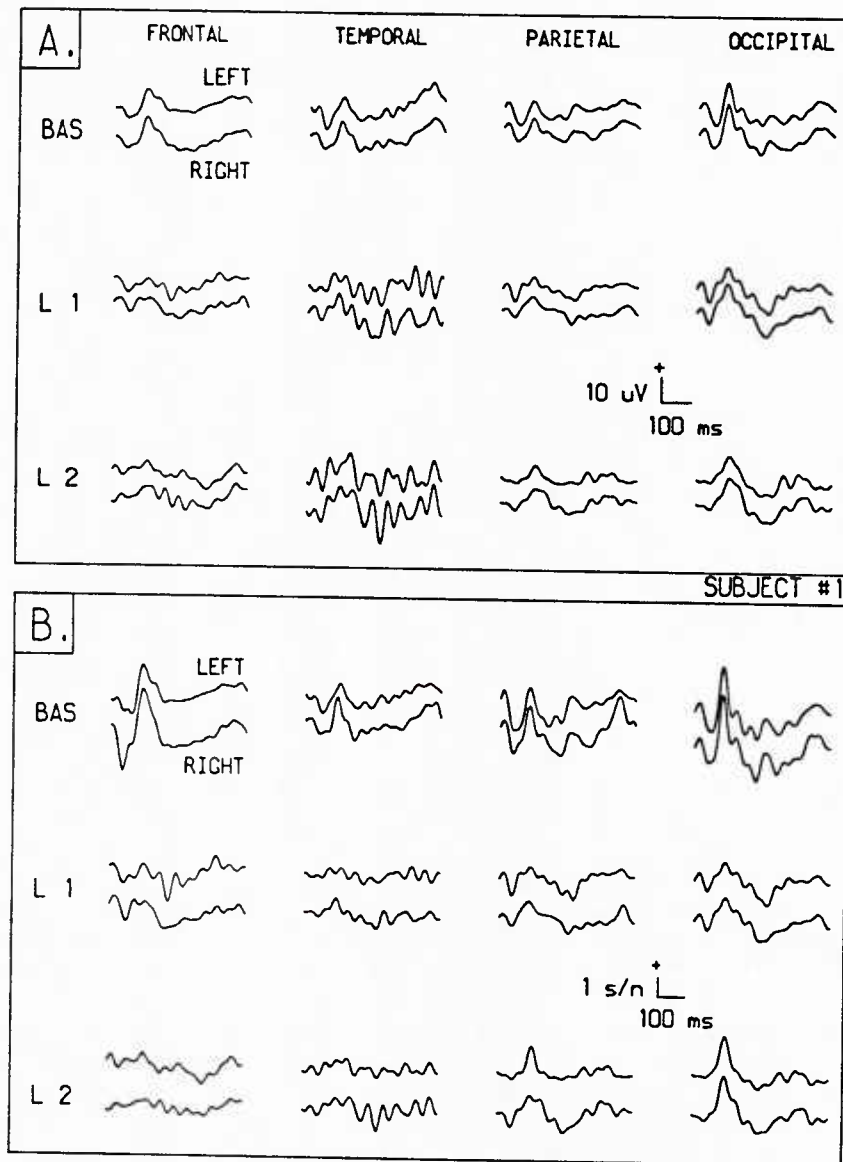


Figure 2. **A.** AWAVES (average EP) and **B.** SNWAVES (signal-to-noise ratio) of one subject are shown for eight recording sites (F3, T3, P3, O1, F4, T4, P4, and O2) across the three experimental conditions (baseline, level 1, and level 2). The ordinate is a millisecond time scale beginning at stimulus onset and extending for a total of 500 ms. The abscissa is a voltage scale for AWAVES and a signal-to-noise ratio scale for SNWAVES. Calibration bars show 100-ms time increments (both measures), a 10 μ V amplitude for AWAVES and a unit signal-to-noise ratio for SNWAVES. Each waveform is computed from six artifact-free, single EPs. The SNWAVE better emphasizes activity that is consistent across trials and better de-emphasizes activity that varies from trial to trial. At several sites, the evoked activity near 100-150 ms shows up as a more prominent peak in the AWAVES than in the SNWAVES. Conversely, the noisy EEG/EMG interference at the temporal sites is more pronounced in the AWAVES than in the SNWAVES. Artifact rejection and computational methods for each measure are described in the text.

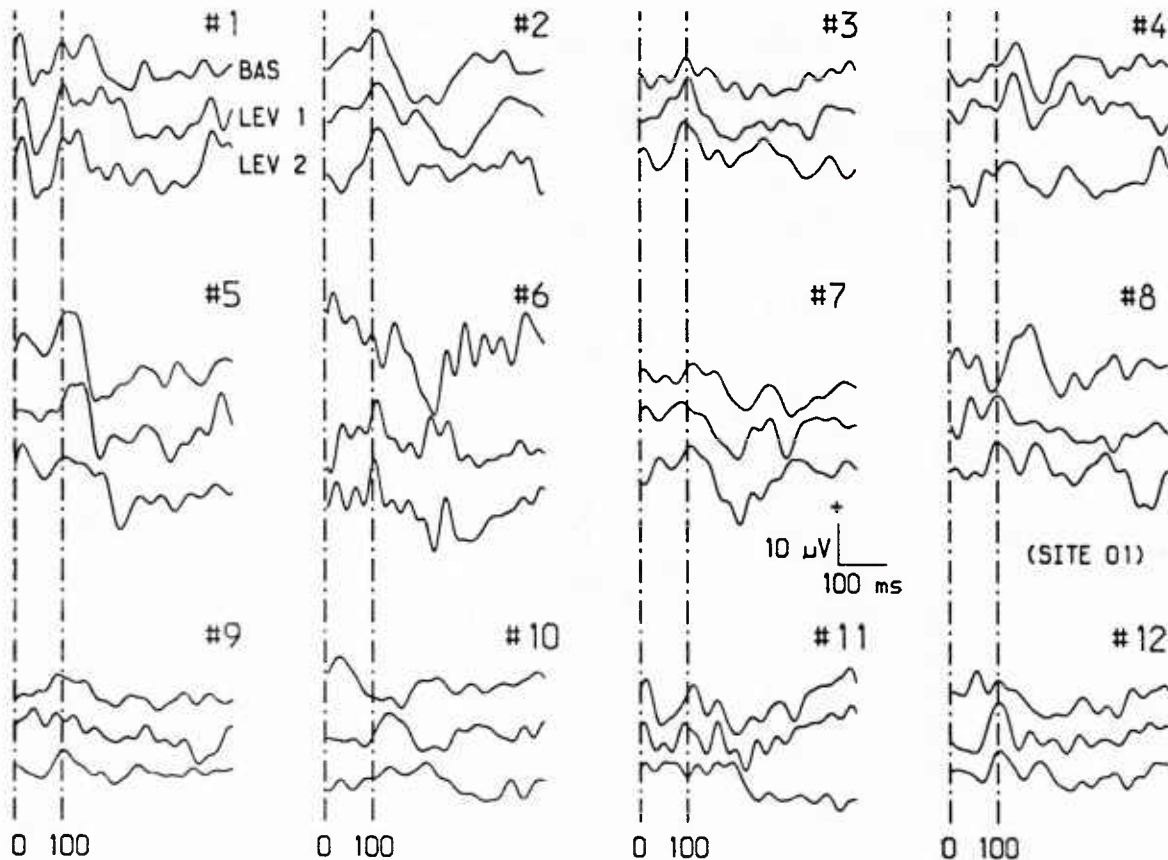


Figure 3. AWAVEs for 12 subjects based on six single, visual-probe EPs recorded over the left occipital area (site O1). Separate traces are plotted for each subject's measurements during the three experimental conditions: baseline (upper trace), level 1 (middle trace), and level 2 (lower trace). EP morphology (waveshape) varies considerably between subjects. Plotting conventions are defined in Figure 2.

the same well-defined EP components, we divided the AWAVE and SNWAVE waveforms into eight non-overlapping latency intervals, each 50.78 ms wide (13 sample points at 256 Hz), spanning the range between 46.88 and 453.13 ms in the post-stimulus period. The beginning, end, and midpoint latencies of each window are listed in Table 2. For each of these windows, we computed an integrated amplitude measure, the unbiased root-mean-square value (RMS) in μV units. These RMS measures served as dependent variables for statistical analyses. The value of the RMS measure for personnel assessment has been discussed elsewhere (Lewis, 1983).

Analyses

A preliminary analysis showed that the typed command and voice command groups did not differ significantly in terms of behavioral or physiological (RMS) measures. Therefore these groups were combined and the analysis was reduced to four independent variables:

- (1) The *workload* factor represents the target frequency in AIRDEF. It has three values corresponding to baseline (no targets), level 1 (4.5 targets per minute), and level 2 (9 targets per minute).
- (2) The *coronal* factor represents one aspect of the position of the recording electrode. As shown in Figure 4, when the scalp is viewed from above, a line running from right to left parallel to the inter-aural line is said to lie in a coronal section or plane. When such lines differ in distance, as measured from the front to the back of the head, they are said to lie in different coronal planes. Two points lying on a single line must lie in a single coronal plane. Thus, the

Table 2

Time Windows for Analyses of EP Waveforms

Window Number	Window Latency (ms)		
	Beginning	End	Midpoint
1	50.78	101.56	76.17
2	101.56	152.34	126.95
3	152.34	203.13	177.74
4	203.13	253.91	228.52
5	253.91	304.69	279.30
6	304.69	355.47	330.08
7	355.47	406.25	380.86
8	406.25	457.03	431.64

coronal factor groups recording sites on opposite sites of the head that lie at the same distance from front to back. The coronal factor has four levels, corresponding to frontal sites (F3, F4), temporal sites (T3, T4), parietal sites (P3, P4), and occipital sites (O1, O2).

- (3) The *sagittal* factor represents another aspect of the position of the recording electrode. As shown in Figure 4, when the scalp is viewed from above, a line running from front to back and perpendicular to the inter-aural line is said to lie in a sagittal section or plane. When such lines differ in distance, as measured from the right side to the left side of the head, they are said to lie in different sagittal planes. Two points lying on a single line must lie in a single sagittal plane. Thus, the sagittal factor groups recording sites that lie at different distances from front to back but lie at the same distance from right to left. The sagittal factor has two levels, corresponding to left hemisphere sites (F3, T3, P3, O1) and right hemisphere sites (F4, T4, P4, O2).
- (4) The *time* factor indexes the AWAVE or SNWAVE RMS values in the different post-stimulus latency windows mentioned previously. The time factor has eight levels corresponding to the eight time windows listed in Table 2.

The dependent variables were *RMS-a* (the RMS amplitude values of the AWAVE in μV units) and *RMS-s* (the RMS values of the SNWAVE in dimensionless units). We performed a separate repeated-measures analysis of variance (Keppel, 1973) for each of the two dependent variables, *RMS-a* and *RMS-s*. Because separate analyses were performed for each dependent measure and because data from different sites and time windows recorded for each subject are often highly correlated, we chose a conservative significance level of $p < .01$ for all effects. Furthermore, we applied to all effects a conservative correction for violations of the assumption of compound symmetry in the computation of the F-ratio (Geisser & Greenhouse, 1958).

RESULTS

Synopsis

Table 3 summarizes the significant main effects and interactions of the four independent variables, workload, coronal electrode position, sagittal electrode position, and time. Separate repeated-measures analyses of variance for the two dependent variables, *RMS-a* (upper half of table) and *RMS-s* (lower half of table) gave similar results. However, *only* *RMS-s*, the signal-to-noise measure, changed significantly as a function of workload. The effect of workload just missed significance for *RMS-a*, the average amplitude measure, $F(2, 58) = 4.65$, $p < .02$. Both *RMS-a* and *RMS-s* differed as functions of coronal (front to back) electrode position and time. In addition, for both measures there was a significant interaction of the coronal and time factors, and a significant three-way interaction of the workload, coronal, and time factors. The sagittal factor was not significant for *RMS-*

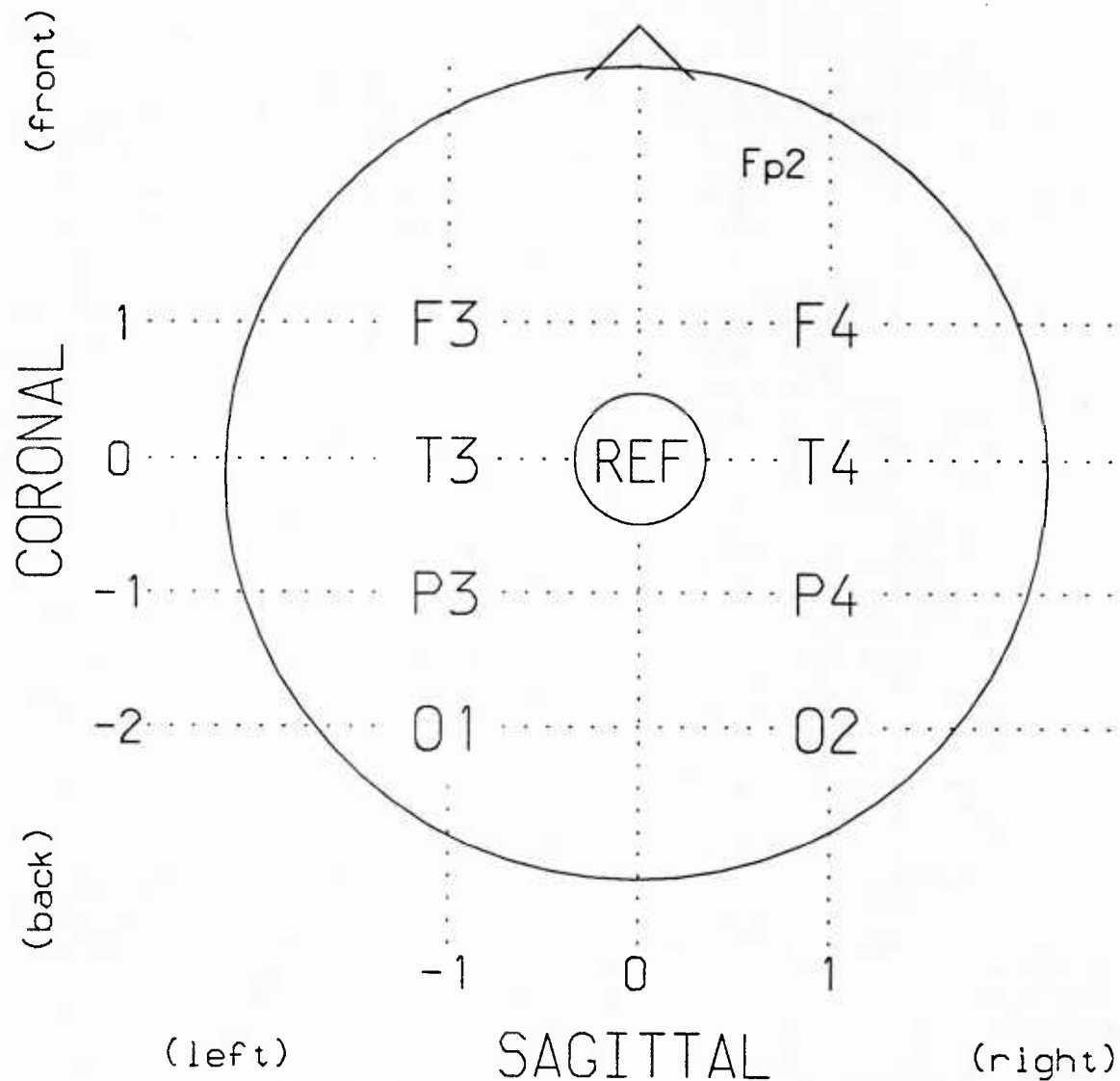


Figure 4. A schematic view of the human head from above. The diagram illustrates the nomenclature used for the electrode recording sites. The vertical line represents the midline; the horizontal line represents the inter-aural line. The eight electrode sites lie in four coronal "planes." (The use of the term *plane* is for convenience and does not imply a strict geometric form.) For example, sites F3 and F4 lie in one coronal plane (+1) anterior to the inter-aural line. The eight sites also lie in two sagittal planes. For example, sites F3, T3, P3, and O1 lie in the left sagittal plane (-1). Coronal positions differ in distance from front to back of the head. Sagittal positions differ in distance from the left side to the right side of head. All recordings were referred to site Cz, which is marked as "REF" in the center of the diagram. Further details are in the text.

Table 3

Analyses of Variance of Evoked Potential RMS-s and RMS-a

<i>A. Root-Mean-Square of Evoked Potential Amplitude</i>			
<i>Source^a</i>	<i>SS</i>	<i>df</i>	<i>F^b</i>
Coronal (C)	1977.14	3	60.23
C × Subjects (S)	952.52	87	
Time (T)	707.03	7	11.70
T × S	1752.63	203	
C × T	473.94	21	6.23
C × T × S	2207.87	609	
Workload (W) × C × T	315.19	42	2.34
W × C × T × S	3902.80	1218	

<i>B. Root-Mean-Square of Evoked Potential Signal-to-Noise Ratio</i>			
<i>Source</i>	<i>SS</i>	<i>df</i>	<i>F^b</i>
Workload (W)	18.19	2	10.97
W × Subjects (S)	48.09	58	
Coronal	11.78	3	22.38
C × S	15.27	87	
Time	11.98	7	10.48
T × S	33.13	203	
C × T	5.30	21	3.76
C × T × S	40.87	609	
W × C × T	6.58	42	2.43
W × C × T × S	78.35	1218	

^aSagittal effects and interactions were not significant, and therefore, are not shown here.

^b $p < .001$ for all these effects.

a, $F(1, 29) = 0.05$, $p < 0.82$, or RMS-s $F(1, 29) = 0.00$, $p < 0.99$.

Effects of Workload

Figure 5, which plots the average RMS-s of homologous sagittal sites, shows that RMS-s was greater during the baseline period than during level 1 and 2 activities. Although not illustrated, this workload-related difference in RMS-s occurred at all eight recording sites. There was a large difference in RMS-s between baseline and level 1, and little or no difference between level 1 and level 2. The mean RMS-s across all sites and time windows was 0.48 for the baseline, 0.35 for level 1 and 0.36 for level 2. If we view this difference as a workload-related decrease in RMS-s, the fractional change between baseline and the average of the two active levels is 25 percent.

Some individuals did not show a decrease in RMS-s as a function of workload. An analysis of these EP differences and their relationship to performance measures is the subject of another report (Trejo, Lewis, & Blankenship, in preparation).

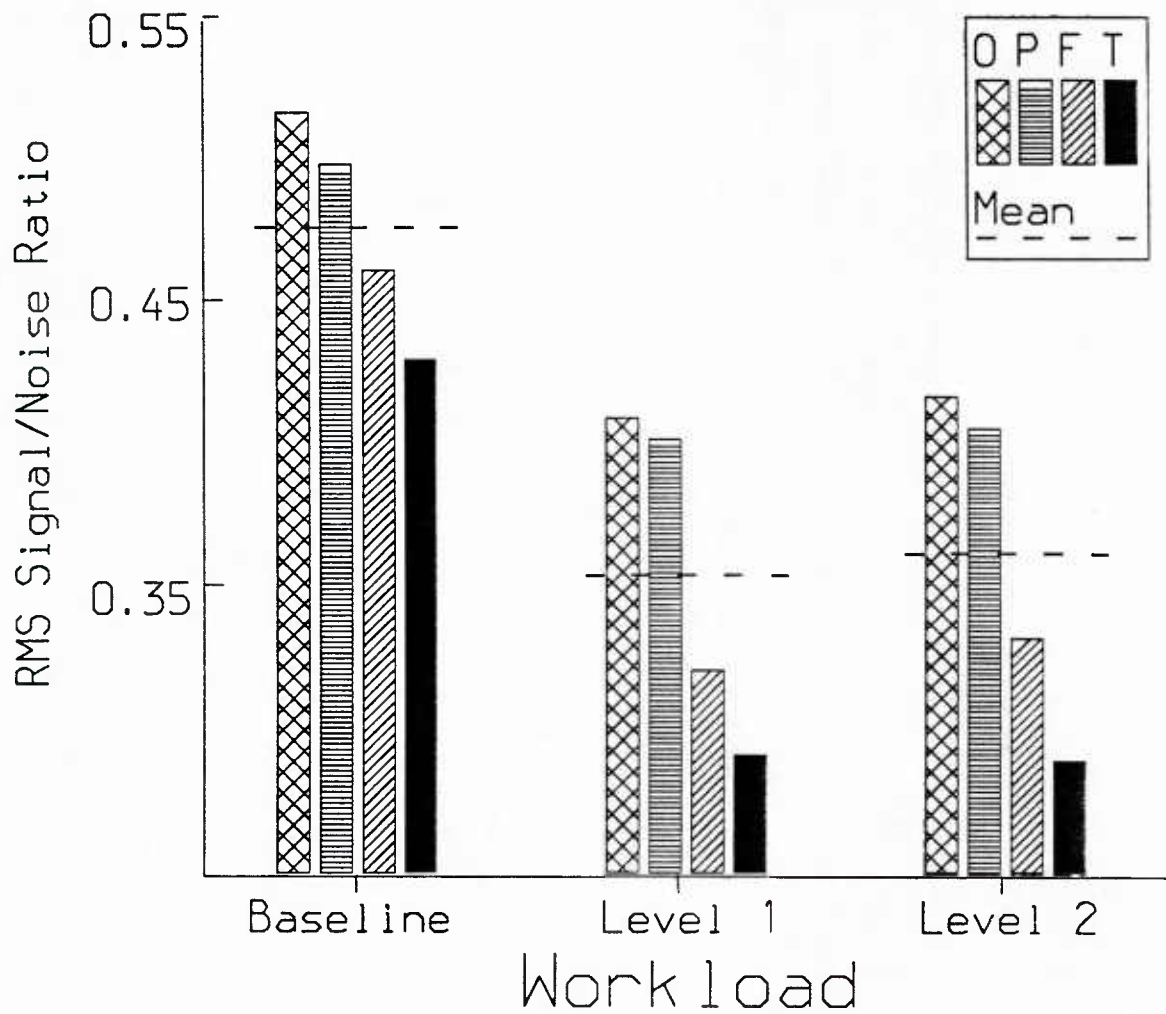


Figure 5. Mean RMS-s (30 subjects) collapsed across all eight time windows (50.78 to 457.03 ms) are plotted as a function of workload. The dashed line shows the mean across all eight recording sites, while the labeled bars show the means for homologous sagittal sites in the four coronal planes: frontal (F), temporal (T), parietal (P), and occipital (O). Mean RMS-s was an average of 25% lower during the two active levels than during the baseline period. Differences between levels 1 and 2 were not significant.

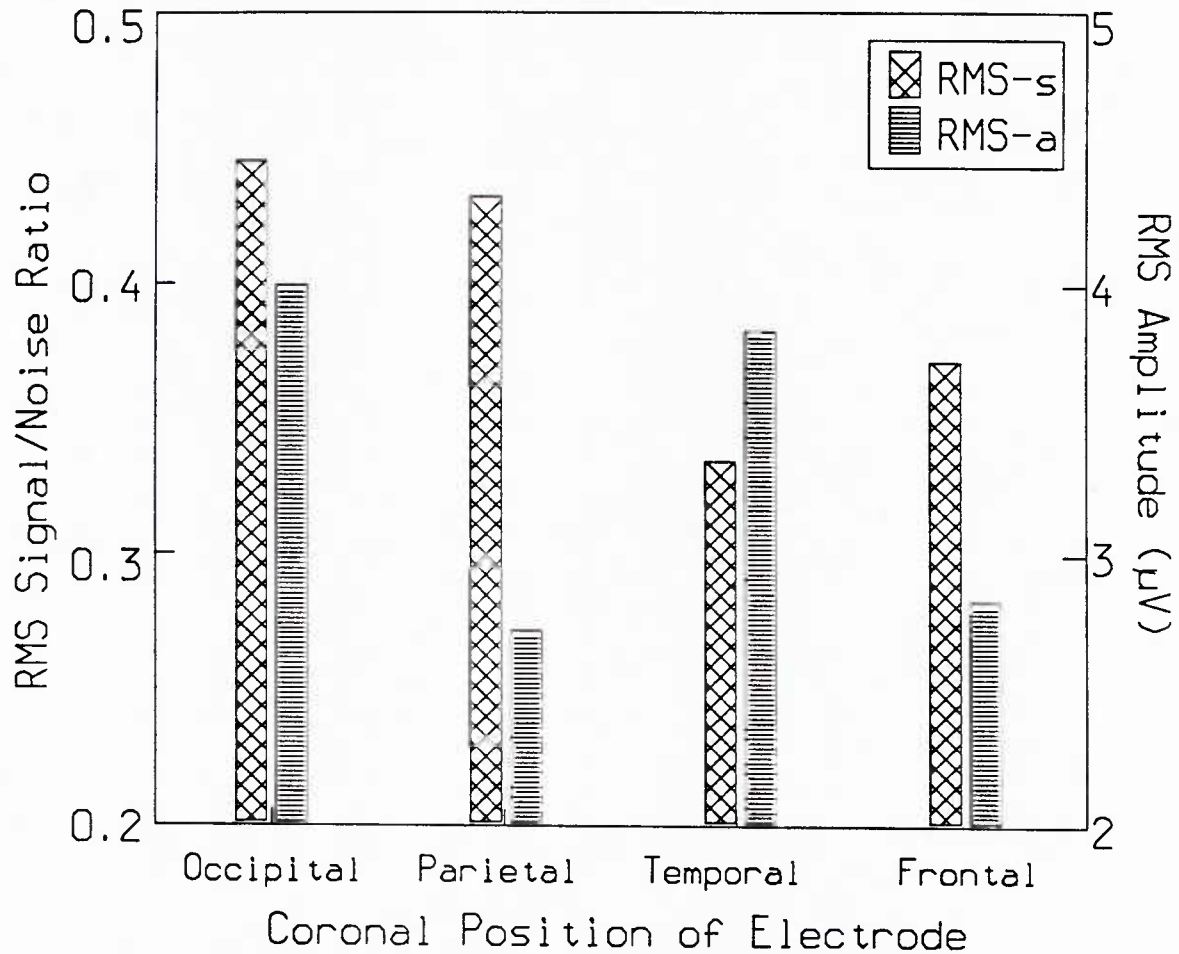


Figure 6. Mean RMS-s and RMS-a (30 subjects) collapsed across time windows and workload levels are shown for each coronal electrode position. Largest values for both measures occurred at the occipital sites. RMS-s was greater at the back of the head (occipital and parietal) than at the front (temporal and frontal). RMS-a values were less orderly. The high RMS-a at temporal sites and low RMS-a at parietal sites suggest that, under our recording conditions, RMS-a is more sensitive to EEG and EMG interference and less sensitive to evoked activity than RMS-s.

Differences Related to Electrode Placement

Mean values of RMS-a and RMS-s for homologous sites in different coronal planes are shown in Figure 6. Both RMS-a and RMS-s were greatest at the occipital sites, as might be expected for a visually evoked potential. For the RMS-a measure, decreasing amplitudes were found for the temporal, frontal, and parietal regions, in that order. For the RMS-s measure, decreasing values were found for the parietal, frontal, and temporal regions, in that order. The rank ordering of the parietal and temporal sites depends strongly on the measure. Implications of these measure-dependent differences for signal extraction using small numbers of single EPs are addressed in the discussion.

Neither RMS-a nor RMS-s depended on the sagittal electrode position (RMS-a: $F(1, 29) = .053$, $p < .82$; RMS-s: $F(1, 29) = .002$, $p < .99$). Thus, on average, left and right brain hemispheres appeared to respond similarly to the probe stimuli under all conditions.

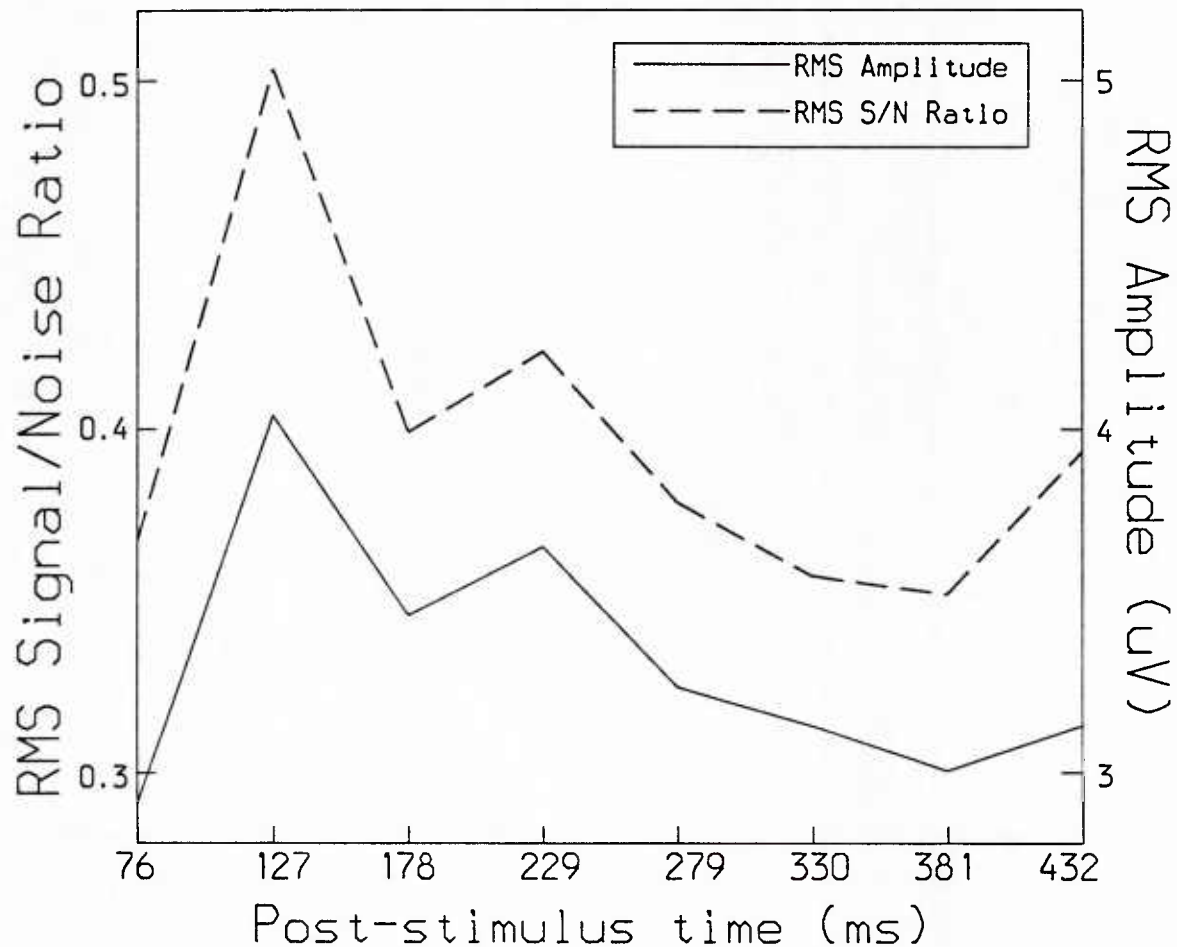


Figure 7. Mean RMS-s and RMS-a (30 subjects) collapsed across the eight recording sites (F3, T3, P3, O1, F4, T4, P4, O2) are plotted as a function of time after stimulus onset. The ordinate indicates the time at the center of each window, as defined in Table 2. The abscissa on the left indicates RMS-s and the abscissa on the right indicates RMS-a in μV units. Both measures show a large maximum in the window centered at 127 ms (window 2) and a smaller maximum in the window centered at 229 ms (window 4).

Differences Related to Post-stimulus Time

Both RMS-a and RMS-s varied across the post-stimulus recording intervals. As Figure 7 shows, both measures showed maxima in windows 2 (127 ms) and 4 (229 ms). Both measures also showed minima in windows 3 (178 ms) and 7 (381 ms).

Interaction of Time and Electrode Placement

For both RMS-a and RMS-s there was a significant interaction of the time and coronal factors. Thus, the strength of the EP signal in different time windows was related to the distance of the electrode from the front to the back of the head. Figure 8 illustrates this interaction for the RMS-s measure. Two major differences are apparent in Figure 8. First, RMS-s is much greater in window 2 (127 ms) at the occipital sites than at all other sites. This peak probably represents early sensory processing of the visual stimulus in the primary visual cortex. Second, the parietal and occipital sites both show a well-defined peak in window 4 (229 ms) that is not obvious at the frontal or temporal sites. It is likely that this peak represents later, higher-order processing of the visual stimulus by secondary

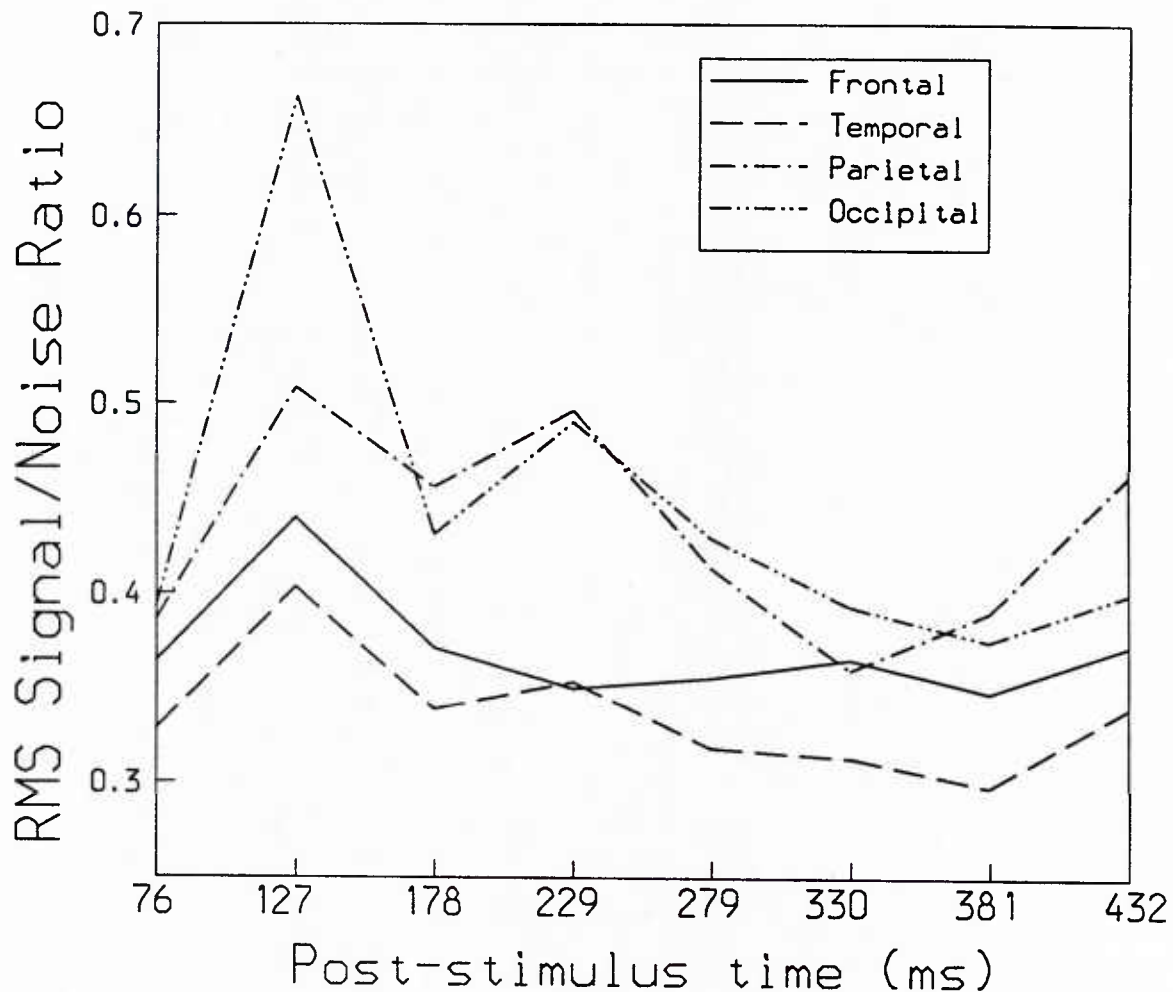


Figure 8. Mean RMS-s (30 subjects) collapsed across homologous recording sites in the same coronal plane (F3 & F4, T3 & T4, P3 & P4, O1 & O2) is plotted as a function of time after stimulus onset. Different lines (see legend) connect the window means of the four coronal electrode positions. The abscissa indicates the time at the center of each window, as defined in Table 2. The ordinate indicates RMS-s. All sites show a maximum in the window centered at 127 ms (window 2), but this maximum is most prominent at the occipital sites. The temporal, parietal, and occipital sites have another maximum in the window centered at 229 ms (window 4). Only the frontal site shows a late maximum, in the window centered at 330 ms (window 6). These curves indicate that the probe stimulus elicits site-specific signals. One possible interpretation is that the activity at the parietal and occipital sites (in windows 2 & 4) represents sensory and early cognitive processing of the visual probe. Higher cognitive processing may be represented by later (window 6) activity at the frontal sites.

visual cortical areas. With the exception of these values, the differences between the remaining curves are not striking. One interesting point is the peak in RMS-s in window 6 (330 ms) at the frontal sites. No other site showed this peak, and evidence to be presented in the next section suggests that it may reflect cognitive processing. For RMS-a (not shown), there was a similar pattern of interaction.

Interaction of Workload, Time, and Electrode Placement

The workload factor interacted with the time and coronal electrode placement factors for both RMS-a and RMS-s. Figure 9 illustrates this interaction. Separate curves in each graph plot the mean values of RMS-s (Figure 9A.) or RMS-a (Figure 9B.) for the four coronal electrode locations. In general, Figure 9 shows that window- and site-specific EP signal reductions occur as workload increases from baseline to active engagement.

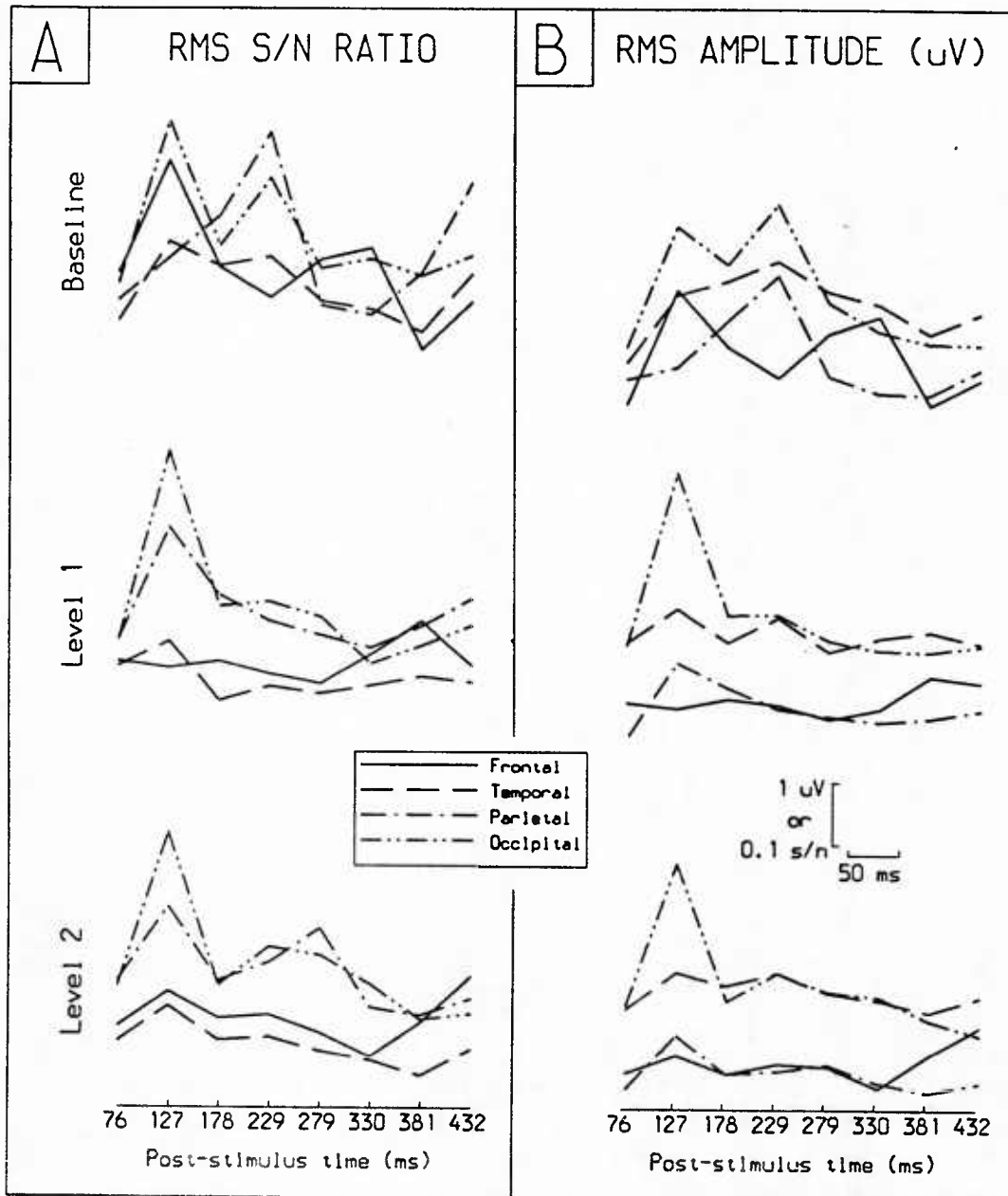


Figure 9. *A.* Mean RMS-s and *B.* RMS-a (30 subjects) collapsed across homologous recording sites in the same coronal plane (F3 & F4, T3 & T4, P3 & P4, O1 & O2) are plotted as a function of time after stimulus onset and separated by workload level. The upper two graphs are for the baseline condition, the middle two for level 1, and the lower two for level 2. As in Figure 8, different lines (see legend) connect the window means of the four coronal electrode positions. The abscissa indicates the time at the center of each window, as defined in Table 2. Scale markers indicate 50-ms time increments and either 1 μ V RMS-a or one unit of RMS-s. Site- and time-dependent workload effects appear to be segregated into two broad categories. First, parietal and occipital RMS-s and RMS-a were greater near 229 ms (window 4) for the baseline condition than for levels 1 and 2. Second, frontal RMS-a and RMS-s for windows centered at 127 ms, 279 ms, and 330 ms (windows 2, 5, and 6, respectively) were also greater for the baseline condition than during levels 1 and 2. The possible significance of these differences is discussed in the text.

The site- and time-specific effects displayed in Figure 9 seem to fall into two categories. First, in window 4 (229 ms), there are workload-related differences in RMS-a and RMS-s at the occipital and parietal sites. We tested the significance of this trend using a general linear model (SAS software, PROC GLM) to compare the mean RMS-s or RMS-a during baseline with the combined mean RMS-s or RMS-a during levels 1 and 2. The results appear in Table 4. The mean values of RMS-s and RMS-a are greater in window 4 (229 ms) during baseline than during the active conditions (levels 1 and 2). One measure of the size of these workload-related differences is the percentage change from baseline to active conditions, which is also shown in Table 4. Larger percentage changes in RMS-s and RMS-a (-41 and -45.8%, respectively) occurred at the parietal sites than at the occipital sites (-29.3 and -29.4%, respectively). The tests confirmed that both RMS-s and RMS-a at frontal sites in the period between 203 and 254 ms after stimulus onset are significantly lower under workload than under baseline conditions.

The second category of effects that appear in Figure 9 is a workload-related difference in RMS-s and RMS-a that is localized to the frontal region and that spans three different time windows. Table 4 shows that frontal RMS-s and RMS-a in windows 2, 5, and 6 were significantly greater during baseline than during the active conditions. Together, these differences span two time intervals: 102 to 152 ms (window 2) and 254 to 355 ms (windows 5 and 6). Percentage change was comparable across the RMS-s and RMS-a measures and across the three time windows, ranging from -33.2 percent to -46.6 percent.

Table 4

Site- and Time-specific Comparisons of Baseline with Active (AIRDEF) EP Measures

<i>A. Comparisons of Baseline RMS-s with Active (AIRDEF) RMS-s</i>							
<i>Time window</i>	<i>Coronal plane</i>	<i>Average RMS-s</i>		<i>Percentage change</i>	<i>SS</i>	<i>F^{a,b}</i>	
		<i>Baseline</i>	<i>AIRDEF</i>				
4	Parietal	0.68	0.40	-41.0	3.13	48.74	
4	Occipital	0.61	0.43	-29.3	1.55	24.04	
2	Frontal	0.64	0.34	-46.6	3.52	54.75	
5	Frontal	0.48	0.29	-38.4	1.34	20.84	
6	Frontal	0.50	0.30	-39.4	1.52	23.66	

<i>B. Comparisons of Baseline RMS-a with Active (AIRDEF) RMS-a</i>							
<i>Time window</i>	<i>Coronal plane</i>	<i>Average RMS-a</i>		<i>Percentage change</i>	<i>SS</i>	<i>F^{a,c}</i>	
		<i>Baseline</i>	<i>AIRDEF</i>				
4	Parietal	4.53	2.46	-45.8	172.49	53.83	
4	Occipital	5.69	4.02	-29.4	162.42	51.32	
2	Frontal	4.29	2.58	-39.9	117.11	36.54	
5	Frontal	3.61	2.41	-33.2	57.19	17.85	
6	Frontal	3.88	2.32	-40.1	96.70	30.18	

^a $p < .001$ for all these effects.

^b $df_{\text{num}} = 1, df_{\text{denom}} = 5664, SS_{\text{error}} = 0.0643.$

^c $df_{\text{num}} = 1, df_{\text{denom}} = 5664, SS_{\text{error}} = 3.2043.$

DISCUSSION

Differences between probe EPs during the baseline and the active periods conformed to the predictions of the resource allocation model and the brain-behavior mapping hypothesis (see Introduction). The resource allocation model predicted that fewer brain resources would be allocated to the probe stimulus during task engagement than during the baseline condition. The mapping hypothesis predicted that the difference in resource allocation between baseline and active engagement would be reflected in the probe-evoked potentials. These predictions were supported by the RMS amplitude and signal-to-noise ratio reductions (Figures 5 and 9) we observed.

Neither RMS-a nor RMS-s differed significantly between the two active engagement levels, suggesting two distinct possibilities. First, task difficulty may not have been effectively greater in level 2 than in level 1. In support of this idea, separate analyses of behavioral performance variables (Trejo, Lewis, & Blankenship, in preparation) were performed. They show that the mean normalized skill rating for AIRDEF (Trejo, 1986) increased slightly, from 36.74 points in level 1 to 44.64 points in level 2, $F(1, 47) = 5.229$, $p < 0.027$. Since subjects had no prior experience with AIRDEF, it is probable that learning during level 1 helped their performance in level 2. Second, larger EP differences between no-workload and workload conditions than those between different workload conditions may be a general property of human brain activity. For example, Israel et al. (1980) reported that EP amplitude between 300 and 500 ms was greater in response to visual probe stimuli in a baseline condition than when subjects monitored display elements, but there was no difference in probe-evoked amplitude between conditions in which either four or eight display elements were monitored. Results like these could be explained by a non-linearity, such as an exponential decay function, in the relationship between probe EP amplitude and workload.

One explanation for the workload-related probe EP reductions we observed is that either the number of neurons that respond to the probe or the intensity of their response decreases in proportion to the resources allocated to a primary task. This view is supported by our finding that probe EPs were lower, on average, during the active engagements than during the baseline condition. RMS-a was an average of 13.9 percent lower and RMS-s 25 percent lower. This view agrees with our general concepts about the sources of neuroelectric activity. If we conceive of processing as the activity of neurons, which are generators of electric currents, then increased processing means either increased numbers of active neurons or increased activity in a number of neurons, which could lead to increased electric currents and to larger evoked potentials.

The alternative to this view is that as processing demands increase, *different* neurons begin to respond to the probe and neurons that responded before cease responding. This view predicts spatial or temporal changes in probe EPs as a function of workload without a necessary increase in the activity or number of active neurons. Some support for this view is given by the significant three-way interaction of the workload, coronal electrode position, and time factors (Figure 9). Furthermore, the lack of significant two-way interactions between workload and site factors or between workload and time factors suggests that workload increases produce *combined* spatial and temporal probe EP changes. As Figure 9 shows, the combined spatial and temporal changes between baseline and active engagement in AIRDEF had two distinct components. First, there was a drop (baseline minus active) in parietal and occipital activity between 200 and 250 ms. Second, there was a drop in frontal activity near 100 and 300 ms.

Two models concerned with the functional significance of EP morphology help to explain the combined spatial and temporal changes in probe EPs that we have observed. First, White, White and Hintze (1984) showed that occipital EP components near 100 and 200 ms probably represent two different levels of visual processing. Activity near 100 ms appears to correspond to primitive sensory input to the striate cortex from the lateral geniculate nucleus and is relatively insensitive to stimulus conditions. Activity near 200 ms appears to correspond to a higher level of sensory processing, probably involving both striate and prestriate cortex, and is very sensitive to stimulus conditions, such

as blur and spatial frequency.⁴ Because early (before 200 ms) parietal and occipital EP signals did not decrease with workload (Figure 9), this view suggests that primitive sensory input to the cortex is not attenuated for probe stimuli as workload increases. However, the drop in activity in window 4 (229 ms) between baseline and active conditions suggests that increasing workload can attenuate the same cortical processes that White et al. (1984) showed to be very sensitive to stimulus conditions. In terms of the resource allocation model, the workload-related effects near 200 ms suggest that task demands may control the flow of sensory input from the geniculo-striate pathway through the striate and prestriate cortex.

The second model of EP morphology and function that helps to explain our workload-related probe EP changes interprets the P300, an endogenous⁴ event-related potential (Sutton, Braren, Zubin, & John, 1965) as an index of mental workload. P300 has been explored in an elegant series of experiments by Donchin and co-workers (reviewed in Gopher & Donchin, 1986). Many other experiments that also bear on the significance of the P300 component for decision-making and workload have been performed (reviewed by Pritchard, 1981), but the essential feature of P300 for our study is its reported ability to index the amount of information or meaning about the environment that is conveyed by a stimulus or event (Sutton, 1979). In our baseline condition, the probe is the only stimulus that the subject views on the display. In the AIRDEF engagements, the probe is one of many stimuli that the subject views, and its environmental relevance is greatly reduced. Thus, if P300 indexes stimulus information or meaning, then we would expect smaller P300 amplitudes or increased P300 latency during levels 1 and 2 than during the baseline condition. Although a P300 in the usual sense was not extracted from our data, the decreases in frontal RMS-s and RMS-a in windows 5 and 6 (250 to 350 ms) (Figure 9, Table 4) support this expectation.

At the frontal electrode sites, there was also a small, but insignificant, increase in RMS-a and RMS-s at windows 7 and 8 (350 to 450 ms) during levels 1 and 2 as compared to the baseline (Figure 9). When accompanied by the decrease in window 6, these increases form a peak in the curves of RMS-s and RMS-a at window 7 during level 1 which is not present during the baseline. During level 2 the peak is no longer in window 7 but instead appears to occur later, in window 8 or beyond. Thus it appears that a peak in the RMS-a and RMS-s curves is displaced toward progressively greater latencies as a function of workload. This is suggestive of a latency shift in late EP components, possibly including P300.

Donchin, Ritter, and McCallum (1978) have proposed that P300 latency is related to the time a subject spends evaluating a stimulus, and that it is independent of any response that he makes. Gopher and Donchin (1986) discussed several experiments that support this proposal. Although P300 effects are typically assessed only with task-relevant stimuli, our data with irrelevant probes may bear on the proposal of Donchin et al. as follows: Because the number of task-relevant stimuli increased in our experiment with workload, subjects probably spent an increasing fraction of the time available for stimulus evaluation on the task-relevant stimuli as workload increased. On the average, this could have progressively delayed the evaluation of the irrelevant probe stimuli as the number of task-relevant stimuli increased. No such delay would have occurred under the baseline condition, where task-relevant stimuli were not presented. This is consistent with our observation of minimal "latency" of a late frontal RMS-s peak during the baseline condition. The maximum delay would have occurred during level 2, and here the "latency" of the late frontal RMS-s peak appears to be greatest.

One effect that is not accounted for by either of the two models discussed above (resource allocation, P300) is the drop in frontal amplitude near 100 ms that we found between baseline and active levels (Figure 9, Table 4). A possible interpretation of this effect is that it reflects a shift in selective attention. There is evidence that a visual EP peak near 150 ms, known as the N1 or N150 component, is sensitive to shifts of attention between different locations (Eason, Harter, & White, 1969; Van Voorhis & Hillyard, 1977). Also, attention-related EP effects have been measured as early as 50

⁴Potentials occurring with latencies below 250 ms are usually called *exogenous* (generated from without) because they occur with regular latency in the presence of a triggering event. *Endogenous* (generated from within) potentials are thought to relate closely to cognitive activity, typically have latencies greater than 250 ms, and are more variable in latency than exogenous potentials. Under some conditions, the brain appears to generate endogenous potentials even in the absence of triggering events.

ms after stimulus onset at frontal sites, and these effects were shown to be reduced or absent in patients who had damage to the frontal lobes (Knight, Hillyard, Woods, & Neville, 1981). We hesitate to conclude that the frontal 100 ms activity changes we measured reflect selective attention, because the existence of pathways that could mediate such rapid attention effects in the frontal lobes remains to be demonstrated.

EP variation and noise may also have contributed to the effect of workload on probe EPs. For example, if EP *latency variation* increases with workload, then workload-related decreases in peak EP amplitudes may be occurring. Such decreases occur because peak latency variation cancels positive and negative single EP peaks when the average EP is computed, thus producing smaller average peak EP amplitudes. Similarly, *amplitude variation* could also depend on workload. Such variation would raise the standard deviation of each point in the average EP while not necessarily reducing its value. This would result in a bigger decrease in a signal-to-noise ratio measure, such as RMS-s, than in a straight average EP measure such as RMS-a. A workload-related increase in amplitude variation is supported by our findings because RMS-s decreased 25 percent between baseline and active levels, whereas RMS-a decreased only 14 percent.

Finally, the eye movements needed for scanning the AIRDEF display during active engagements are also a potential source of some latency and amplitude variation. Eye movements can change the effectiveness of a visual stimulus by projecting it on different retinal areas, which can vary in sensitivity and response time. However, the likelihood that eye movements account for much of the amplitude effects we have observed is small because the probe stimulus covered the full screen of the display, extending well beyond the boundaries of the AIRDEF simulation. As long as the subject's eyes viewed the display, a large retinal area was stimulated.

One striking result of this study was that workload-related amplitude changes were similar across widely separated brain recording sites. Although absolute amplitude and signal-to-noise ratios differed from the front to the back of the head (Figure 6), different brain recording sites generally showed the same pattern of workload-related changes in both RMS-a and RMS-s (Figure 5). This finding is also supported by the lack of a significant interaction between either of the electrode position factors (coronal or sagittal) and the workload factor. For recordings from all the sites, both RMS-a and RMS-s declined between baseline and the active levels, but did not change between the active levels. An economic implication in this particular study is that a single active electrode could suffice to provide data for assessing workload effects on EP amplitudes during complex decision-making tasks. The use of the vertex (Cz) as a reference for our recordings may also have contributed to the widespread nature of the workload-related EP changes we observed. If brain activity generated near the reference electrode was strongly related to workload, then the reference would not have been a truly indifferent electrode, and widespread workload-related EP effects would have been measured at all sites using this reference.

CONCLUSIONS

- (1) Irrelevant probe EPs reflected differences between baseline and workload (levels 1 and 2) conditions but not between different workload conditions. Across time and electrode sites, root-mean-square EP signal-to-noise ratio was smaller during task performance than during the baseline condition. These findings suggest that a simple integrated amplitude measure from a single scalp electrode could provide an indicator of combat system workload when irrelevant probe stimuli are used. The failure of probe EP amplitudes to discriminate between levels 1 and 2 could be due to non-linearity in the relation between workload and probe EP amplitude. Alternately, workload differences between levels 1 and 2 may have been offset by learning during the level 1 performance, which effectively reduced the *subjective workload* of level 2.
- (2) Workload-related differences in the probe-EP signal-to-noise ratio were about two times larger than amplitude differences. This finding suggests greater EP amplitude and latency variation under workload conditions than under baseline conditions. Furthermore, this result illustrates the need to explore a range of dependent measures and signal transformations in attempts to relate EP measures to complex task conditions.

- (3) The most prominent peak in RMS-a and RMS-s EP measures, which occurred at the back of the head 100-150 ms after stimulus onset, appeared to be unaffected by workload manipulations. Smaller peaks, near 200-250 ms at the back of the head and near 100-150 ms and 300-350 ms at the front of the head, were smaller under workload than under baseline conditions. These results suggest that workload manipulations in a complex combat system simulation affect high-level processing of sensory input from irrelevant probes but not early sensory processing. An implication of this result in terms of resource allocation models of brain processing is that allocation "policy" in the brain may be enforced by controlling the ability of sensory neurons to access cognitive processing mechanisms. In essence, this type of resource allocation is predicted by theories of selective attention that propose higher order mechanisms that select the inputs to be attended to by inhibiting or blocking the inputs to be ignored. (Other theories of selective attention have proposed facilitation or "gating" of the inputs to be attended.)
- (4) The apparent amplitude differences in the 250-350 ms frontal RMS-a and RMS-a may have been caused by a workload-related latency increase rather than a simple amplitude reduction. Such latency increases have been found in other studies for the P300 EP component with task-relevant stimuli, where it appears that the latency of P300 serves as an index of subjects' stimulus evaluation time. We discussed how increases in the number of task-relevant stimuli during simulation performances could have delayed evaluation of our irrelevant probe stimuli, thus producing a latency increase in frontal EP activity near 300-350 ms.

RECOMMENDATIONS

- (1) The AIRDEF task combined with the irrelevant probe EP technique forms a useful tool for studying human information processing in combat system simulations. However, two experimental design changes are necessary in order to discriminate between competing explanations (non-linearity or learning effects) of the equality of amplitude and signal-to-noise ratio in levels 1 and 2. First, a wider range of task difficulty levels must be examined in order to determine whether a non-linear relationship exists between workload and probe EP amplitude. Second, the use of different difficulty levels should be counterbalanced to control for effects of time on task, presentation order, and learning.⁵ If, under these conditions, there are still little or no differences in probe-evoked potential measures between widely disparate difficulty levels, then we may conclude that the relationship between workload and irrelevant probe EP amplitudes is non-linear.
- (2) In the present study, the differences between probe EPs during the baseline and active engagement periods may have been due in part to factors other than decision-making. For example, eye movements and motor responses (key presses and vocal commands) were required for active engagements but not for the baseline condition. The baseline condition should be modified so that it mimics all of the major elements of the active levels, while still requiring no decision-making.⁵ In this way, any probe-evoked potential differences may be seen to relate more clearly to decision-making.
- (3) Combat system operators are often faced with performing secondary tasks using another sensory modality. They may, for example, be required to monitor a radar and attend to verbal instructions. For this reason, the real-world relevance of AIRDEF should be increased by adding a secondary task that requires another sensory modality. Potentials evoked by stimuli used for the secondary task (e.g., counting auditory tones of different frequencies) performed concurrently with AIRDEF, should be explored.⁵

⁵These changes have already been implemented in a revision of this experiment, and data have been acquired in 100 subjects (Trejo, Blankenship, & Lewis, in preparation).

REFERENCES

- Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.
- Callaway, E. (1975). *Brain electrical potentials and individual psychological differences*. New York: Grune & Stratton.
- Donchin, E., Ritter, W., & McCallum, C. (1978). Cognitive psychophysiology: The endogenous components of the ERP. In E. Callaway, P. Tueting, and S. Koslow (Eds.), *Brain event related potentials in man*. New York: Academic Press.
- Eason, R., Harter, M., & White, C. (1969). Effects of attention and arousal on visually evoked cortical potentials and reaction time in man. *Physiology and Behavior*, **4**, 283-289.
- Geisser, S., & Greenhouse, S. W. (1958). An extension of Box's results on the use of the *F* distribution in multivariate analysis. *Annals of Mathematical Statistics*, **29**, 885-891.
- Gevens, A. S. (1986). *Neuro-electric predictors of performance accuracy* (AFOSR Rep. 85-0361). Washington, DC: Air Force Office of Scientific Research.
- Gopher, D., & Donchin, E. (1986). Workload — an examination of the concept. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of perception and human performance*. Vol. II. *Cognitive processes and performance*. New York: John Wiley.
- Israel, J. B., Wickens, C. D., Chesney, G. L., & Donchin, E. (1980). The event-related potential as an index of display-monitoring workload. *Human Factors*, **22**, 211-224.
- Jasper, H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, **10**, 371-375.
- John, E. R., Prichep, L., Ahn, H., Easton, P., Fridman, J., & Kaye, H. (1983). Neurometric evaluation of cognitive dysfunctions and neurological disorders in children. *Progress in Neurophysiology*, **21**, 239-290.
- Karis, D., Coles, M. G. H., & Donchin, E. (1984). *The use of event-related brain potentials for prediction of performance quality in pilots of advanced tactical jet aircraft* (Tech. Rep. CPL 84-2/F33615-82-K-0611). Champaign: Cognitive Psychophysiology Laboratory, University of Illinois.
- Kelly, R. T., Greitzer, F. L., & Hershman, R. L. (July 1981). *Air Defense: A computer game for research in human performance* (NPRDC Tech. Rep. 81-15). San Diego: Navy Personnel Research and Development Center.
- Keppel, G. (1973). *Design and analysis*. Englewood Cliffs, NJ: Prentice Hall.
- Knight, R. T., Hillyard, S. A., Woods, D. L., & Neville, H. J. (1981). The effects of frontal cortex lesions on event-related potentials during auditory selective attention. *Electroencephalography and Clinical Neurophysiology*, **52**, 571-52.
- Kramer, A. F., Wickens, C. D., & Donchin, E. (1983). An analysis of the processing requirements of a complex perceptual-motor task. *Human Factors*, **25**, 597-621.

- Lewis, G. W. (1983). Event related brain electrical and magnetic activity: Toward predicting on-job performance. *International Journal of Neuroscience*, 18, 159-182.
- Lewis, G. W., & Rimland, B. (1979). *Hemispheric asymmetry as related to pilot and radar intercept officer performance* (NPRDC Tech. Rep. 79-13). San Diego: Navy Personnel Research and Development Center.
- Lewis, G. W., & Rimland, B. (1980). *Psychobiological measures as predictors of sonar operator performance* (NPRDC Tech. Rep. 80-26). San Diego: Navy Personnel Research and Development Center.
- Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica*, 27, 84-92.
- Norman, D., & Bobrow, D. (1975). On data limited and resource limited processing. *Journal of Cognitive Psychology*, 7, 44-60.
- Papanicolaou, A. C., & Johnstone, J. (1984). Probe evoked potentials: Theory, method and applications. *International Journal of Neuroscience*, 24, 107-131.
- Pritchard, W. S. (1981). Psychophysiology of P300. *Psychological Bulletin*, 89, 506-50.
- Regan, D. (1972). *Evoked potentials in psychology, sensory physiology, and clinical medicine*. London: Chapman & Hall.
- Sutton, S. (1979). P300 — 13 years later. In H. Begleiter (Ed.), *Evoked brain potentials and behavior*. New York: Plenum Press.
- Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked potential correlates of stimulus uncertainty. *Science*, 150, 1187-1188.
- Taylor, M. M., Lindsay, P. M., & Forbes, S. M., (1967). Quantification of shared capacity processing in auditory and visual discrimination. *Acta Psychologica*, 27, 223-229.
- Trejo, L. J. (1986). *Brain activity during tactical decision-making: I. Hypotheses and experimental design* (HFOSL Tech. Note 71-86-6). San Diego: Navy Personnel Research and Development Center.
- Trejo, L. J., Blankenship, M. H., & Lewis, G. W., (In preparation). *Probe-evoked potential amplitude measures and decision-making performance under varying workload: Effects of stimulus modality and a secondary task*.
- Trejo, L. J., Lewis, G. W., & Blankenship, M. H. (In preparation). *Brain activity during tactical decision-making: III. Relationships between physiology, simulation performance, and job performance*.
- Van Voorhis, S. T., & Hillyard, S. A. (1977). Visual evoked potentials and selective attention to points in space. *Perception and Psychophysics*, 22, 54-62.

- White, C. T., White, C. L., & Hintze, R. W. (1984). Functional differences of early and late components of pattern-reversal visual evoked potential. In J. R. Heckenlively, G. H. M. Van Lith, and T. Lawwill (Eds.), *Documenta Ophthalmologica Proceedings Series*, 40. Boston: Dr. W. Junk Publishers.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman and D. R. Davies (Eds.), *Varieties of attention*. New York: Academic Press.

DEPARTMENT OF THE NAVY
NAVY PERSONNEL RESEARCH AND
DEVELOPMENT CENTER
(CODE _____)
SAN DIEGO, CA 92152-6800
OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, \$300

U232845